# RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: _/0/632,793A_
Source: _1Fw/6_
Date Processed by STIC: _10/26/06_

# *ENTERED*

RAW SEQUENCE LISTING         DATE: 10/26/2006
PATENT APPLICATION:  **US/10/632,793A**    TIME: 10:36:18

        Input Set : **A:\11004801.txt**
        Output Set: **N:\CRF4\10262006\J632793A.raw**

```
 3 <110> APPLICANT: PARANHOS-BACCALA, Glaucia
 4        MALLET, Francois
 5        VOISSET, Cecile
 7 <120> TITLE OF INVENTION: ENDOGENEOUS NUCLEIC ACID FRAGMENT ASSOCIATED WITH AN
AUTOIMMUNE
 8        DISEASE, LABELING METHOD AND REAGENT
10 <130> FILE REFERENCE: 110048.01
12 <140> CURRENT APPLICATION NUMBER: 10/632,793A
13 <141> CURRENT FILING DATE: 2003-08-04
15 <150> PRIOR APPLICATION NUMBER: PCT/FR00/00144
16 <151> PRIOR FILING DATE: 2000-01-21
18 <150> PRIOR APPLICATION NUMBER: US 09/869,927
19 <151> PRIOR FILING DATE: 2001-08-17
21 <150> PRIOR APPLICATION NUMBER: FR 99/00888
22 <151> PRIOR FILING DATE: 1999-01-21
24 <160> NUMBER OF SEQ ID NOS: 36
26 <170> SOFTWARE: PatentIn version 3.3
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 1511
30 <212> TYPE: DNA
31 <213> ORGANISM: Homo sapiens
33 <400> SEQUENCE: 1
34 cctagaacgt attctggaga attgggacca atgtgacact cagacgctaa gaaagaaacg       60
36 atttatattc ttctgcagta ccgcctggcc acaatatcct cttcaaggga gagaaacctg      120
38 gcttcctgag ggaagtataa attataacat catcttacag ctagacctct tctgtagaaa      180
40 ggagggcaaa tggagtgaag tgccatatgt gcaaactttc ttttcattaa gagacaactc      240
42 acaattatgt aaaaagtgtg gtttatgccc tacaggaagc cctcagagtc cacctcccta      300
44 ccccagcgtc ccctccccga ctccttcctc aactaataag gacccccctt taacccaaac      360
46 ggtccaaaag gagatagaca aaggggtaaa caatgaacca aagagtgcca atattccccg      420
48 attatgcccc ctccaagcag tgagaggagg agaattcggc ccagccagag tgcctgtacc      480
50 tttttctctc tcagacttaa agcaaattaa aatagaccta ggtaaattct cagataaccc      540
52 tgacggctat attgatgttt tacaagggtt aggacaatcc tttgatctga catggagaga      600
54 tataatgtta ctactaaatc agacactaac cccaaatgag agaagtgccg ctgtaactgc      660
56 agcccgagag tttggcgatc tttggtatct cagtcaggcc aacaatagga tgacaacaga      720
58 ggaaagaaca actcccacag gccagcaggc agttcccagt gtagacctc attgggacac      780
60 agaatcagaa catggagatt ggtgccacaa acatttgcta acttgcgtgc tagaaggact      840
62 gaggaaaact aggaagaagc ctatgaatta ctcaatgatg tccactataa cacagggaaa      900
64 ggaagaaaat cttactgctt ttctggacag actaagggag gcattgagga agcatacctc      960
66 cctgtcacct gactctattg aaggccaact aatcttaaag gataagttta tcactcagtc     1020
68 agctgcagac attagaaaaa acttcaaaag tctgccttag gcccggagca gaacttagaa     1080
70 accctattta acttggcatc ctcagttttt tataatagag atcaggagga gcaggcgaaa     1140
72 cgggacaaac gggataaaaa aaaaggggg ggtccactac tttagtcatg gccctcaggc     1200
74 aagcagactt tggaggctct gcaaaaggga aaagctgggc aaatcaaatg cctaataggg     1260
76 ctggcttcca gtgcggtcta caaggacact ttaaaaaaga ttatccaagt agaaataagc     1320
```

RAW SEQUENCE LISTING                    DATE: 10/26/2006
PATENT APPLICATION:  US/10/632,793A     TIME: 10:36:18

        Input Set : **A:\11004801.txt**
        Output Set: **N:\CRF4\10262006\J632793A.raw**

```
 78 cgcccccttg tccatgcccc ttacgtcaag ggaatcactg gaaggcccac tgccccaggg     1380
 80 gatgaagata ctctgagtca gaagccatta accagatgat ccagcagcag gactgagggt     1440
 82 gcccggggcg agcgccagcc catgccatca ccctcacaga gccccgggta tgtttgacca     1500
 84 ttgagagcca a                                                          1511
 87 <210> SEQ ID NO: 2
 88 <211> LENGTH: 2009
 89 <212> TYPE: DNA
 90 <213> ORGANISM: Homo sapiens
 93 <220> FEATURE:
 94 <221> NAME/KEY: misc_feature
 95 <222> LOCATION: (1940)..(1940)
 96 <223> OTHER INFORMATION: n = a or g or c or t/u
 98 <400> SEQUENCE: 2
 99 atacgactac tatagggcga attgggccct ctagatgcat gctcgagcgg ccgccagtgt       60
101 gatggatatc tgcagaattc gcccctttgtc cgctgtgctc ctgatccagc gaggcgccca      120
103 ttgctgctcc caattgggct aaaggcttgc cattgttccc acacggctaa gtgcccgggt      180
105 tcatcctaat tgagctgaac actagtcact gggttccatg gttctcttcc atgacccacg      240
107 gcttctaata gagctctaat actcaccaca tggcccaaga ttccattcct tggaatccgt      300
109 gaggccaaga accccaggtc agagaacacg aggcttgcca ccgtcttgga agtggcccgc      360
111 cgccatcttg ggagctctgg gagcaaggac cccccagtaa cattttggca accacaaagg      420
113 gacctccaaa gcgatgggaa acattccccc caaggcaaaa acgcccctaa gatgtattct      480
115 ggagaattgg gaccaatgtg acactcagac gctaagaaag aaacgattta tattcttctg      540
117 cagtaccgcc tggccacaat atcctcttca agggagagaa acctggcttc ctgagggaag      600
119 tataaattat aacatcatct tacagctaga cctcttctgt agaaaggagg gcaaatggag      660
121 tgaagtgcca tatgtgcaaa cttctctttc attaagagac aactcacaat tatgtaaaaa      720
123 gtgtggttta tgccctacag gaagccctca gagtccacct ccctacccca gcgtcccccc      780
125 cccgactcct tcctcaacta ataaggaccc cccttttaacc caaacggtcc aaaaggagat      840
127 agacaaaggg gtaaacaatg aaccaaagag tgccaatatt ccccgattat gccccctcca      900
129 agcagtgaga ggaggagaat tcggcccagc cagagtgcct gtaccttttt ctctctcaga      960
131 cttaaagcaa attaaaatag acctaggtaa attctcagat aaccctgacg gctatattga     1020
133 tgttttacaa gggttaggac aatcctttga tctgacatgg agagatataa tgttactact     1080
135 aaatcagaca ctaaccccaa atgagagaag tgccgctgta actgcagccc gagagtttgg     1140
137 cgatctttgg tatctcagtc aggtcaacaa taggatgaca acagaggaaa gaacaactcc     1200
139 cacaggccag caggcagttc ccagtgtaga ccctcattgg gacacagaat cagaacatgg     1260
141 agattggtgc cacaaacatt tgctaacttg cgtgctagaa ggactgagga aaactaggaa     1320
143 gaagcctatg aattactcaa tgatgtccac tataacacag ggaaaggaag aaaatcctac     1380
145 tgcttttctg gacagactaa gggaggcatt gaggaagcat acctccctgt cacctgactc     1440
147 tattgaaggc caactaatct taaaggataa gtttatcact cagtcagctg cagacattag     1500
149 aaaaaacttc aaaagtctgc cttaggcccg gagcagaact tagaaaccct atttaacttg     1560
151 gcatcctcag ttttttataa tagagatcag gaggagcagg cgaaacggga caaacgggat     1620
153 aaaaaaaaaa gggggggtcc actactttag tcatggccct caggcaagca gactttggag     1680
155 gctctggaaa agggaaaagc tgggcaaatc aaatgcctaa tagggctggc ttccagtgcg     1740
157 gtctacaagg acactttaaa aaagattatc caagtagaaa taagccgccc ccttgtccat     1800
159 gccccttacg tcaagggaat cactggaagg cccactgccc caggggatga agatactctg     1860
161 agtcagaagc cattaaccag atgatccagc agcaggactg agggtgcccg gggcgagcgc     1920
W--> 163 **cagcccatgc catcaccctn acagagcccc gggtatgctt gaccattgag agccaggagg**     **1980**
165 ttaactgtct cctggacact ggcgcagcc                                        2009
168 <210> SEQ ID NO: 3
169 <211> LENGTH: 1056
```

RAW SEQUENCE LISTING                    DATE: 10/26/2006
PATENT APPLICATION:  US/10/632,793A       TIME: 10:36:18

Input Set : **A:\11004801.txt**
Output Set: **N:\CRF4\10262006\J632793A.raw**

```
170 <212> TYPE: DNA
171 <213> ORGANISM: Homo sapiens
173 <400> SEQUENCE: 3
174 ctagaacgta ttctggagaa ttgggaccaa tgtgacactc agacgctaag aaagaaacga     60
176 tttatattct tctgcagtac cgcctggcca caatatcctc ttcaagggag agaaacctgg    120
178 cttcctgagg gaagtataaa ttataacatc atcttacagc tagacctctt ctgtagaaag    180
180 gagggcaaat ggagtgaagt gccatatgtg caaactttct tttcattaag agacaactca    240
182 caattatgta aaaagtgtgg tttatgccct acaggaagcc ctcagagtcc acctccctac    300
184 cccagcgtcc cctccccgac tccttcctca actaataagg accccccttt aacccaaacg    360
186 gtccaaaagg agatagacaa aggggtaaac aatgaaccaa agagtgccaa tattccccga    420
188 ttatgccccc tccaagcagt gagaggagga gaattcggcc cagccagagt gcctgtacct    480
190 ttttctctct cagacttaaa gcaaattaaa atagacctag gtaaattctc agataaccct    540
192 gacggctata ttgatgtttt acaagggtta ggacaatcct ttgatctgac atggagagat    600
194 ataatgttac tactaaatca gacactaacc ccaaatgaga gaagtgccgc tgtaactgca    660
196 gcccgagagt ttggcgatct ttggtatctc agtcaggcca acaataggat gacaacagag    720
198 gaaagaacaa ctcccacagg ccagcaggca gttcccagtg tagaccctca ttgggacaca    780
200 gaatcagaac atggagattg gtgccacaaa catttgctaa cttgcgtgct agaaggactg    840
202 aggaaaacta ggaagaagcc tatgaattac tcaatgatgt ccactataac acagggaaag    900
204 gaagaaaatc ttactgcttt tctggacaga ctaagggagg cattgaggaa gcatacctcc    960
206 ctgtcacctg actctattga aggccaacta atcttaaagg ataagtttat cactcagtca   1020
208 gctgcagaca ttagaaaaaa cttcaaaagt ctgcct.                            1056
211 <210> SEQ ID NO: 4
212 <211> LENGTH: 25
213 <212> TYPE: DNA
214 <213> ORGANISM: Homo sapiens
216 <400> SEQUENCE: 4
217 cggacatcca aagtgatggg aaacg                                           25
220 <210> SEQ ID NO: 5
221 <211> LENGTH: 26
222 <212> TYPE: DNA
223 <213> ORGANISM: Homo sapiens
225 <400> SEQUENCE: 5
226 ggacaggaaa gtaagactga gaaggc                                          26
229 <210> SEQ ID NO: 6
230 <211> LENGTH: 26
231 <212> TYPE: DNA
232 <213> ORGANISM: Homo sapiens
234 <400> SEQUENCE: 6
235 cctagaacgt attctggaga attggg                                          26
238 <210> SEQ ID NO: 7
239 <211> LENGTH: 26
240 <212> TYPE: DNA
241 <213> ORGANISM: Homo sapiens
243 <400> SEQUENCE: 7
244 tggctctcaa tggtcaaaca tacccg                                          26
247 <210> SEQ ID NO: 8
248 <211> LENGTH: 21
249 <212> TYPE: DNA
250 <213> ORGANISM: Homo sapiens
```

RAW SEQUENCE LISTING          DATE: 10/26/2006
PATENT APPLICATION:  US/10/632,793A     TIME: 10:36:18

Input Set : A:\11004801.txt
Output Set: N:\CRF4\10262006\J632793A.raw

252 <400> SEQUENCE: 8
253 atgggaaacg ttccccccga g                                      21
256 <210> SEQ ID NO: 9
257 <211> LENGTH: 21
258 <212> TYPE: DNA
259 <213> ORGANISM: Homo sapiens
261 <400> SEQUENCE: 9
262 ggcctaaggc agacttttga a                                      21
265 <210> SEQ ID NO: 10
266 <211> LENGTH: 409
267 <212> TYPE: PRT
268 <213> ORGANISM: Homo sapiens
270 <400> SEQUENCE: 10
272 Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val Pro
273 1               5                   10                  15
276 Arg Gly Ser His Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg
277             20                  25                  30
280 Ile Met Gly Asn Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile
281         35                  40                  45
284 Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg
285     50                  55                  60
288 Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly
289 65                  70                  75                  80
292 Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu
293                 85                  90                  95
296 Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro
297             100                 105                 110
300 Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys
301         115                 120                 125
304 Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr
305     130                 135                 140
308 Pro Ser Val Pro Pro Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro
309 145                 150                 155                 160
312 Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu
313                 165                 170                 175
316 Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg
317             180                 185                 190
320 Gly Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser
321         195                 200                 205
324 Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro
325     210                 215                 220
328 Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu
329 225                 230                 235                 240
332 Thr Trp Arg Asp Ile Met Leu Leu Leu Asn Gln Thr Leu Thr Pro Asn
333                 245                 250                 255
336 Glu Arg Ser Ala Ala Val Thr Ala Ala Arg Glu Phe Gly Asp Leu Trp
337             260                 265                 270
340 Tyr Leu Ser Gln Val Asn Asn Arg Met Thr Thr Glu Glu Arg Thr Thr
341         275                 280                 285

```
344 Pro Thr Gly Gln Gln Ala Val Pro Ser Val Asp Pro His Trp Asp Thr
345     290             295             300
348 Glu Ser Glu His Gly Asp Trp Cys His Lys His Leu Leu Thr Cys Val
349 305             310             315             320
352 Leu Glu Gly Leu Arg Lys Thr Arg Lys Lys Pro Met Asn Tyr Ser Met
353             325             330             335
356 Met Ser Thr Ile Thr Gln Gly Lys Glu Glu Asn Pro Thr Ala Phe Leu
357             340             345             350
360 Asp Arg Leu Arg Glu Ala Leu Arg Lys His Thr Ser Leu Ser Pro Asp
361             355             360             365
364 Ser Ile Glu Gly Gln Leu Ile Leu Lys Asp Lys Phe Ile Thr Gln Ser
365     370             375             380
368 Ala Ala Asp Ile Arg Lys Asn Phe Lys Ser Leu Pro Lys Leu Ala Ala
369 385             390             395             400
372 Ala Leu Glu His His His His His His
373             405
376 <210> SEQ ID NO: 11
377 <211> LENGTH: 393
378 <212> TYPE: PRT
379 <213> ORGANISM: Homo sapiens
381 <400> SEQUENCE: 11
383 Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Ile Met Gly Asn
384 1               5               10              15
387 Ile Pro Pro Lys Ala Lys Thr Pro Leu Arg Cys Ile Leu Glu Arg Ile
388             20              25              30
391 Leu Glu Asn Trp Asp Gln Cys Asp Thr Gln Thr Leu Arg Lys Lys Arg
392         35              40              45
395 Phe Ile Phe Phe Cys Ser Thr Ala Trp Pro Gln Tyr Pro Leu Gln Gly
396     50              55              60
399 Arg Glu Thr Trp Leu Pro Glu Gly Ser Ile Asn Tyr Asn Ile Ile Leu
400 65              70              75              80
403 Gln Leu Asp Leu Phe Cys Arg Lys Glu Gly Lys Trp Ser Glu Val Pro
404             85              90              95
407 Tyr Val Gln Thr Phe Phe Ser Leu Arg Asp Asn Ser Gln Leu Cys Lys
408             100             105             110
411 Lys Cys Gly Leu Cys Pro Thr Gly Ser Pro Gln Ser Pro Pro Pro Tyr
412     115             120             125
415 Pro Ser Val Pro Ser Pro Thr Pro Ser Ser Thr Asn Lys Asp Pro Pro
416     130             135             140
419 Leu Thr Gln Thr Val Gln Lys Glu Ile Asp Lys Gly Val Asn Asn Glu
420 145             150             155             160
423 Pro Lys Ser Ala Asn Ile Pro Arg Leu Cys Pro Leu Gln Ala Val Arg
424             165             170             175
427 Gly Gly Glu Phe Gly Pro Ala Arg Val Pro Val Pro Phe Ser Leu Ser
428             180             185             190
431 Asp Leu Lys Gln Ile Lys Ile Asp Leu Gly Lys Phe Ser Asp Asn Pro
432     195             200             205
435 Asp Gly Tyr Ile Asp Val Leu Gln Gly Leu Gly Gln Ser Phe Asp Leu
436     210             215             220
```

RAW SEQUENCE LISTING ERROR SUMMARY       DATE: 10/26/2006
PATENT APPLICATION:  **US/10/632,793A**       TIME: 10:36:19

*FYI*

Input Set : **A:\11004801.txt**
Output Set: **N:\CRF4\10262006\J632793A.raw**

**Please Note:**
**Use of n and/or Xaa have been detected in the Sequence Listing. Please review the**
**Sequence Listing to ensure that a corresponding explanation is presented in the <220>**
**to <223> fields of each sequence which presents at least one n or Xaa.**

Seq#:2; N Pos. 1940
Seq#:18; N Pos. 594
Seq#:22; N Pos. 879,1200
Seq#:23; N Pos. 305
Seq#:24; N Pos. 84,193,241
Seq#:27; N Pos. 119
Seq#:29; N Pos. 1191,1213,2089,2274
Seq#:30; N Pos. 198,307,355,1309,1331,2213,2398,3787,4115,4261

**VERIFICATION SUMMARY**                    DATE: 10/26/2006
PATENT APPLICATION:  **US/10/632,793A**        TIME: 10:36:19

Input Set : **A:\11004801.txt**
Output Set: **N:\CRF4\10262006\J632793A.raw**

```
L:163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:1920
L:567 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:540
L:804 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:840
M:341 Repeated in SeqNo=22
L:847 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:300
L:930 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:60
M:341 Repeated in SeqNo=24
L:1155 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:60
L:1364 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29 after pos.:1140
M:341 Repeated in SeqNo=29
L:1470 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:180
M:341 Repeated in SeqNo=30
```